TIME-FREQUENCY DECOMPOSITION BY MEANS OF PHYSICAL MODEL

Eric INCERTI, Claude CADOZ
email : Eric.Incerti@imag.fr Claude.Cadoz@imag.fr

A.C.R.O.E.
(Association pour la Création et la Recherche sur les Outils d'Expression)
I.N.P.G.
Institut IMAG
(Informatique et Mathématiques Appliquées de Grenoble)
(Unité Fédérative 0001 CNRS)
46 Av. Felix Viallet 38031 Grenoble - France

ABSTRACT

Physical model is the raw material for many simulation systems. For example, the CORDIS-ANIMA system, designed for sound synthesis. Like every synthesis system, it has dual properties for sound analysis. We presents here a tool, based on the physical model, and specially designed for the analysis of sounds on a time/frequency domain. This is done by means of a decomposition on a basis made of causal elementary mechanical oscillators.

1. INTRODUCTION.

Sound Synthesis by physical model developed at the ACROE with the CORDIS-ANIMA (C.A.) system [1],[2] is based on the simulation of vibrating structures from their elementary mechanical components (matter points, visco-elastic links) represented in the digital world by some specific algorithms that describe the fundamental laws of Newton's mechanics. These elements are the raw material for the construction of a wide range of complex network-based vibrating structures for sound synthesis and animated images production.
These same simulated mechanical components can be used for the realisation of tools, integrated in the C.A. universe, and designed for the analysis (in different ways) of objects created and musical sounds produced by the system.
This paper presents, after a review of the C.A. formalism, a tool designed for spectral analysis (frequency decomposition) of digital sounds.
Let us precise that this contribution should not be seen in a signal processing framework : we stern on a description of properties and capabilities of the physical model as a synthesis and analysis system for musical sound. Anyway we shall pay great attention to the study of links with standard signal processing tools.

2. PHYSICAL MODEL

2.1 Physical Phenomena.

Physical phenomena are produced by the evolution of some system. Their representation is usually backed up on a description of state equations that govern this evolution. The handled abstractions may have no direct real correspondent (for example, the Fourrier Transform of a signal). The formal resolution of these state equations can lead to an explicit representation of the phenomenon.
Then, computers allow a simulation of these phenomena by a direct discretisation of formal solutions. This can also be done, like in the C.A. context, by a direct computation of successive states from previous states and general equations. Hence, the evolution of the system is represented by a sequence of discrete states instead of an analytic formula.

2.2 CORDIS-ANIMA formalism

This formalism enables the computer-based modelisation and simulation of physical objects that can be seen, heard and handled. An object is an assembly of elementary mechanical components picked up among a limited number of types with very simple associated algorithms. The simulation of the global object results from the combination of these algorithms. This combination may be extremely complex. It may generate phenomena that would be very difficult to describe in an analytical way.
In the C.A. formalism, physical interactions are represented by input-output pairs called *communication points*. There are two kind of such points:
- the first one, named `M Point`, receives an intensive variable (force) in input, and returns an extensive variable (position) in output.

- the other one, named `L Point`, receives a position and returns a force.

The construction of a C.A. object, is the assembly of some components with communication points. These points are connected, according to the connection law : `M Point` on `L Point`.

The smallest C.A. components are of two dual types *(fig.1)* :

- The **<MAT>** atoms, named *Material Point*, that have a spatial location and represent the matter granulation; their algorithms return a position from the different forces applied by their environnement.
- The **<LIA>** atoms or *Link Elements*, with two associated `L Points`. Their only spatial location is given by the two `M Points` on which they are connected. Their algorithms return two forces from the two positions they receives. The fundamental principle of Action-Reaction says that these two forces are co-axial, equal in intensity, and in opposite directions.

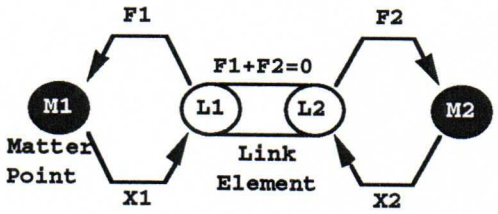


Fig. 1 : Atoms and connections

Among these atoms, three are very important and used for any vibrating structure :

- The free mass : made of a single matter point with a mass parameter `m` (kg). Its algorithm is :

$$(1) \quad X_{n+1} = 2X_n - X_{n-2} + \frac{1}{mT_e^2} F_{n-1}$$

With $T_e = 1/F_e$ (F_e is the computation sample rate)

- The visco-elastic link : link element with a stiffness parameter `k` (N/m) and a damping parameter `z` (Ns/m).

$$(2) \quad F_1 = (k + zT_e)\left(X_n^2 - X_n^1\right) - zT_e\left(X_{n-1}^2 - X_{n-1}^1\right)$$

$$\text{and} \quad F_2 = -F_1$$

- The ground, or fixed point : absorbing matter point. It returns a constant position whatever it receives. Its only parameter is its position. Actually it represents a matter point with an infinite mass.

2.4 Synthesis and Analysis.

This formalism and the associated algorithms allows the simulation of a great variety of physical phenomena. The network of matter points and visco-elastic links can be excited in different ways. Under definite parametric conditions, the nework comes back to its equilibrium state with a damped vibrating motion. Thus, with given initial conditions, the simulation is done from a description of the object by decomposition on a state equations basis. The output information is as a physical states sequence describing a temporal evolution. Each state is a vector representing the values, at a time, of some physical magnitudes. This sequence of numerical states constitutes the synthesis signal. In a practical situation, only a part of this information will be picked up and interpreted. About this point, Flandrin gives (in [3]) an interesting definition :

"Signal is a physical medium of an information (...); it is generally the time evolution of a magnitude picked up at the output of one or several sensors. It makes the raw material from which the 'signal processor' try to extract the information that he considers to be useful" [1].

For example, in the case of a simulation for sound synthesis with C.A., this information will be the position of a particular matter point of the object, interpreted as a sound signal (via a D.A.C., to a loudspeaker). In such a case, the notion of sound frequency comes directly from the notion of eigen frequency of the simulated mechanical network. This parallel between the components of a sound and the eigen elements of the underlaying network gives a new way to define the difficult notion of frequency.

If this definition is clear for a periodic infinite signal [3] (definition used in Fourrier analysis), things are not so simple for real complex and time-limited signals :

- In one hand, there is no sense to deal with frequency without considering the time dimension (this lays in the fundamental mathematic definition of frequency, that is the inverse of a time). This can be seen as a microscopic aspect of time in frequency.
- On the other hand, and specially in the case of sound analysis, the frequency usually decreases with time, or it can be modulated. More generally, frequency includes all the physical parameters of the excitation mode, transcients evolution ...This would be a macroscopic aspect.

The elementary mechanical oscillator (named *Cell* in the C.A. terminology) enables us to take into account this double time aspect (a third aspect

[1] Translated from French by the author.

appears in the sample rate F_e in case of discrete signal). This cell, that constitutes the basic element for any simulation is described on*(fig.2)*.

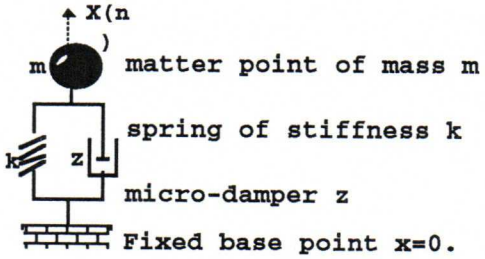


fig.2 : Elementary mechanical oscillator

The evolution of this vibrating element is caracterised by the discrete relation :

$$(3)\quad X_{n+1} = \left(2 - \frac{K+Z}{M}\right)X_n + \left(\frac{Z}{M} - 1\right)X_{n-1}$$

where $K = k$, $Z = zT_e$ and $M = mT_e^2$

With this state equation, a formal resolution gives a description of the damped vibrating motion :

$$(4)\quad X_n = \chi \rho^n \cos(n\omega + \varphi)$$

where the *eigen elements* are :

$$(5)\quad \rho = \sqrt{1 - \frac{Z}{M}} \quad \text{and} \quad \cos(\omega) = \frac{2M - K - Z}{2\sqrt{M^2 - ZM}}$$

χ and φ depending on initial conditions

Remark : m (kg), k (N/m), z (Ns/m) are called physical parameters and M, K, Z, are called algorithmic parameters (considered as pure scalars).

With this representation, these two aspects of time can be separated. The parameter (K/M) can be associated to the microscopic aspect [2], and the parameter (Z/M) can be associated to the macroscopic damping aspect of time. Of course, the exact expression of the eigen frequency of a cell *(6)* brings us back to the Gabor-Heisenberg inequality [3]. It precises that a signal cannot be concentrated on time and frequency domains that would be arbitrarily narrow.
Thus in one hand, this mechanical representation of vibrating phenomenon shows frequency as an intensive dynamic parameter of signal. On the other hand it enables a decomposition of this phenomenon on a causal basis *(fig.6)*. Modal synthesis [4], [5] shows that any linear vibrating network can be decomposed into cells, each cell adjusted on a mode (eigen value) of the network. The signal emitted by the network can be obtained by adding the elementary damped sinusoïdal signals of each cell (excitation conditions being also decomposed on this basis).
The same idea can be used for the decomposition of a signal issued from a network. By dispatching this signal on a bank of cells correctly adjusted, each cell will act like a band pass filter for its eigen frequency and therefore it gives means to detect the presence and the evolution of this frequency in the signal.
Of course, classical signal processing techniques enable to achieve such an analysis, but it seems interesting to use the same philosophy of physical model for construction, simulation, synthesis on one hand, and analysis on the other hand.
Moreover, the same decomposition can be used for the analysis of any kind of sound by considering it as a signal made of sampled positions, issued from an hypothetic simulated network. This will be discussed in the following section.

[2] Frequency of the conservative oscillator (z=0) is :

$$(6)\quad \cos(\omega_{max}) = 1 - \frac{K}{2M}$$

3. FREQUENCY DECOMPOSITON

3.1 Principle.

The elementary oscillator used for the analysis of digital sounds of any origine is a specially designed version of the first one *(fig.2)* . It needs the use of a special ground type element (see section 2.2.) : the vibrating base point. Its position follows the vibrating motion read in the signal to be analysed.This could be seen as a little mass fixed by a spring on a loud-speaker. This system is represented by the standard state equation *(3)*, modified to take into account the vibrating base.
It can be shown that this expression offers a degree of freedom that may be neglected. This enables us to consider every matter point with an algorithmic mass `M=1`. So the system is entirely caracterised by two parameters `K` and `Z`.
Then, the square modulus of the transfer function (or amplitude response) and eigen elements of the system depend only on these two parameters :

$$(7)\quad |H(\omega)|^2 = f(K, Z)$$

The resonance pulse (*eigen pulse*) is given by :

$$(8)\quad \cos(\omega_{max}) \underset{Z\ small}{\approx} 1 - \frac{K}{2}$$

The approximation made in *(8)* comes down to consider the eigen pulse of the system with $z=0$ (this reduced expression is the same that in *(6)*). This approximation is essential for the clarity of analytic expressions because of the complexity of exact relations. For example, with this reduced expression of ω_{max}, the magnitude response is :

$$(9) \quad |H(\omega_{max})|^2 = \frac{K}{Z^2}$$

A bank of such oscillators (ie. band pass filters) can be built by duplicating this elementary system. Each cell is adjusted on a given resonance frequency and damping time through the algorithmic parameters K^i and Z^i .
Various physical properties can be attributed to this bank of cells. This is done by the determination of special relations between the raws of parameters $(K^i)_{i<n}$ and $(Z^i)_{i<n}$ (where n is the number of cells). For example it can be designed for a constant magnitude response :

$$(10) \quad |H^i(\omega^i_{max})|^2 = \alpha \xrightarrow{\text{with (9)}} Z^i = \alpha\sqrt{K^i}$$

Then, the cells basis is build from a given range of frequencies (f_i) and algorithmic parameters given by *(8)* for the K^i and by *(10)* for the Z^i.

(fig.4) shows such amplitude response for three of these cells of respective resonance frequency of 1000, 2000 and 3000 Hz with α=0.2. Values of the eigen and band pass frequency are calculated with the complete analytical expressions. The effect of the aproximation made in *(8)* can be noticed on *(fig.4)* (error rate is less than 1%).

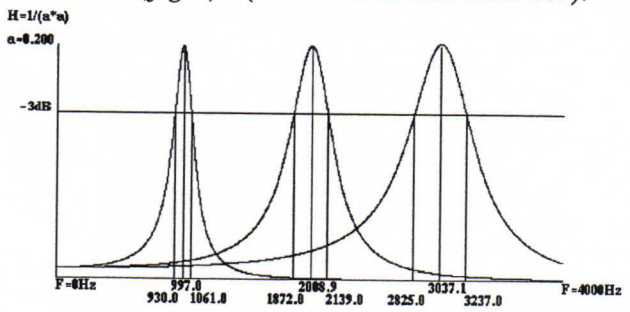


fig.4 : Magnitude response $|H^i(\omega^i_{max})|^2 = \alpha = 0,2$

(fig.5) shows the amplitude response and *(fig.6)* shows the impulse response of a cell for different values of α. This coefficient acts on the damping parameter of each cell and on the bandwidth of associated band pass filter. So, this coefficient α (for a bank with constant magnitude response) represents the Gabor-Heisenberg inequality : a small value of α gives very frequency selective cells (narrow bandwidth) but with high damping time. So it provides good accuracy for the microscopic time aspect but not for the macroscopic time aspect (see 2.4.).

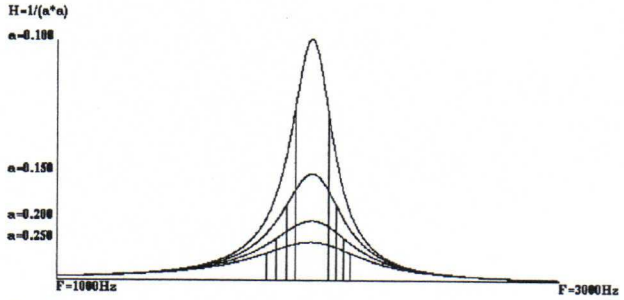


fig.5 : Magnitude response for different α *and f=2000Hz.*

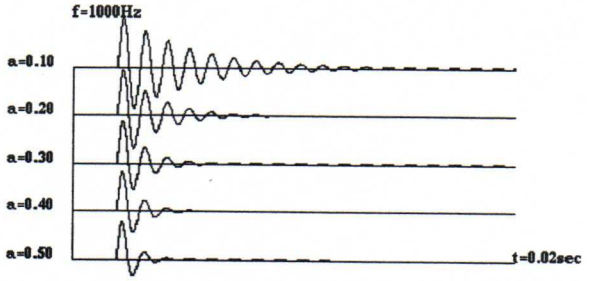


fig.6 : Impulse response for different α *and f=1000Hz.*

But, it is also possible to build a bank of cells with other properties, such as constant bandwidth, constant energy response...what the 'signal processor' wants to point out.

Another very interesting property of such a decomposition basis, is that the impulse response *(fig.6)* of a cell is causal : it does not introduce any anticipation on the signal as with classical wavelets or Fourrier decomposition [6], [7].
Moreover, the fact that each cell is totally independant from each others, avoids interferency phenomena *(fig.10)*, due to an analytical coupling between decomposition elements in standard analytical techniques such as Wigner-Ville ([3] p.238). This also allows a great freedom in the choice of the decomposition basis. For example it is possible to proceed to an analysis of a sound between 0 and 1000Hz with a 50Hz linear scale (20 cells) together with an analysis between 1000 and 10000Hz with a logarithmic scale.
When all the parameters are correctly adjusted, the execution is done always in the same way, by computation of successive states for every cell, each sate corresponding to a sample element of the signal to be analysed.

Of course the computation time of the algorithm is based only on the number of cells and is the same for every cell. For each cell and each sample element of the signal, the algorithm needs only 3 multiplications, 4 additions and 3 data transfers.

3.2 Results

The only information that this system returns is a bank of wavelet-like elementary signals as shown in *(fig.7)* . The interpretation of these signals can be done in many different ways.

One of the simplest ways provides a graphic time/frequency representation *(fig.9, 10, 11)* :

- each elementary signal is cut up with a time window ΔT (ΔT=0,05 sec in *(fig.9)*).
- For each section of time W(i) of each signal S(j), the maximum magnitude A(i,j) is extracted and a new plot is drawn at (W(i), S(j), A(i,j)).

In the examples shown, time domains are separated but it is possible to use slidings windows for more accuracy.

Moreover, in these examples, the time window is the same for every signal (remind that each signal corresponds to a specific frequency). One more time it is possible to use a specific ΔT for each frequency. For example $\Delta T_i=(N/f_i)\times F_e$. That correspond to a time window adjusted on N periods for every frequency. This guarantees an equivalent response delay for every cell.

In the same way, any kind of representation can be done : logarithmic scales, representation of energy, phasis...according to what the user is looking for.

So, the decomposition of a complex signal into a set of independant elementary signals allows great freedom for the interpretation and representation *(fig.8b)*. Classical signal processing techniques usually make a decomposition of the time-frequency space ($s(t)=\sum_t\sum_f T_f(t)\Phi_t(f)$) before processing *(fig.8a)* . This introduces interference elements and strong relations between time and frequency domains. The decomposition on a bank of cells avoids both correlations and provides great flexibility in processing and interpretation.

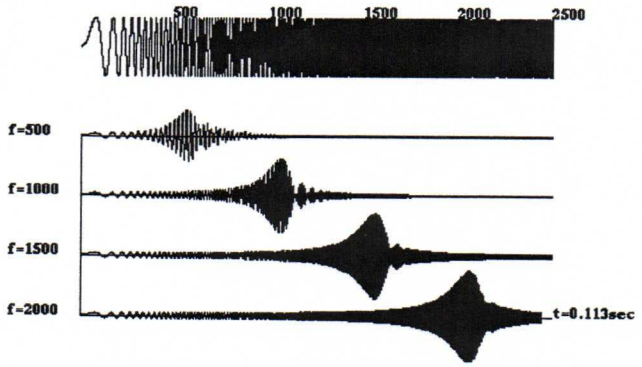
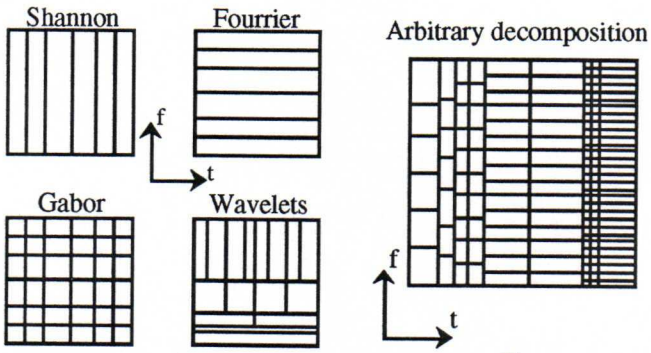


fig.7 : Input signal : linear chirp from 0 to 2000Hz



(a) : *fig. 2.1.* p.55 of [3] (b)

fig.8 Time-Frequency decompositions

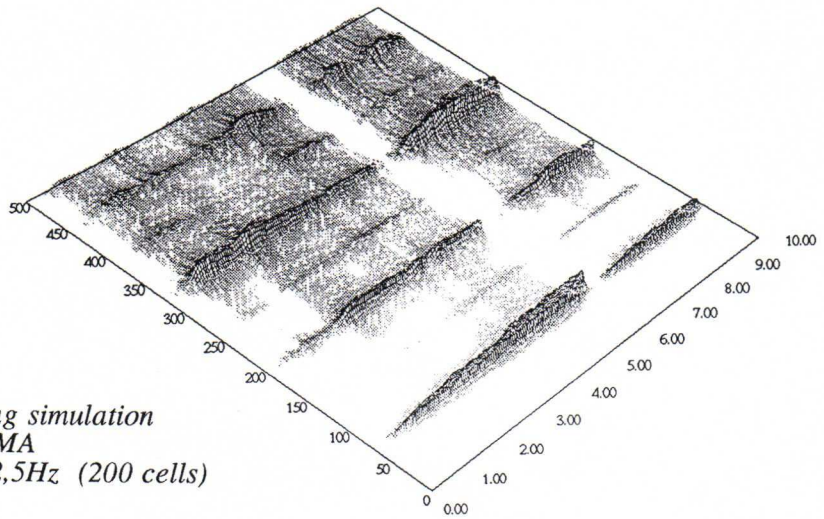


fig.9 : bowed string simulation with CORDIS-ANIMA
ΔT=0,05sec ΔF=2,5Hz (200 cells)

4. CONCLUSION

Thus, it appears that physical model principles can lead to very simple and natural techniques for synthesis (ie. simulation of physical phenomena) but also for analysis (interpretation of these phenomena). The tool described here was designed with special goal : try to caracterise the presence and the evolution of frequency components in a sound signal. We already get promissing results. Now it seems interesting to study with great attention the links between such algorithms and classical signal processing techniques. Moreover, it appears that this particular tools takes his place in a wider context of synthesis and analysis. Some other little tools already exist (selective amplificators, modal caracterisation of vibrating structures...). Physical model allows that : synthesis, then analysis, with the same theoretical elements.

References

1. Cadoz, C., Luciani,A., Florens J.L., "CORDIS-ANIMA : A Modeling and Simulation System for Sound and Image Synthesis - The General Formalism." CMJ Vol.17, N°4, pp19-29 - Spring1993.

2. Flandrin,P, *Temps-Fréquence.* Hermes 1993.

3. Djoharian, P., "Generating Models for Modal Synthesis" CMJ Vol.17, N°4, pp57-65 - Spring1993.

4. Ewins, D.J., *Modal Testing : Theory and Practice.* Research Study Press, 1984

5. Meyer,Y., Jaffard,S., Rioul,O., "L'analyse par ondelettes." Pour la Science. Sept. 1987.

6. Kronland-Martinet,R., "The Wavelet Transform for Analysis, Synthesis and Processing of Speech and Music Sound.", CMJ Vol.12, N°4 pp 11-20 - Winter 1988.

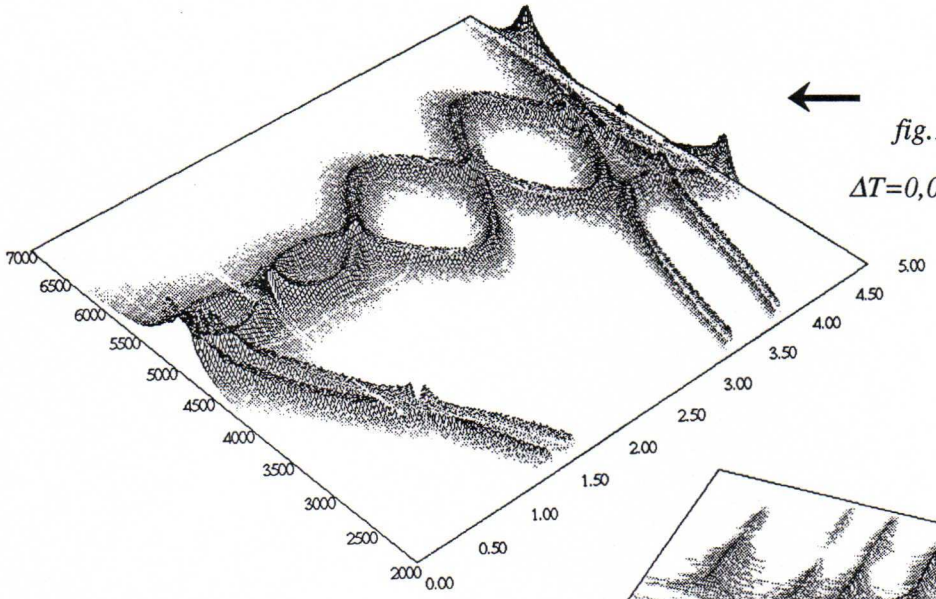


fig.10 : 'exotic' modulated chirp (multiplicative synthesis) ΔT=0,025sec ΔF=25Hz (200 cells)

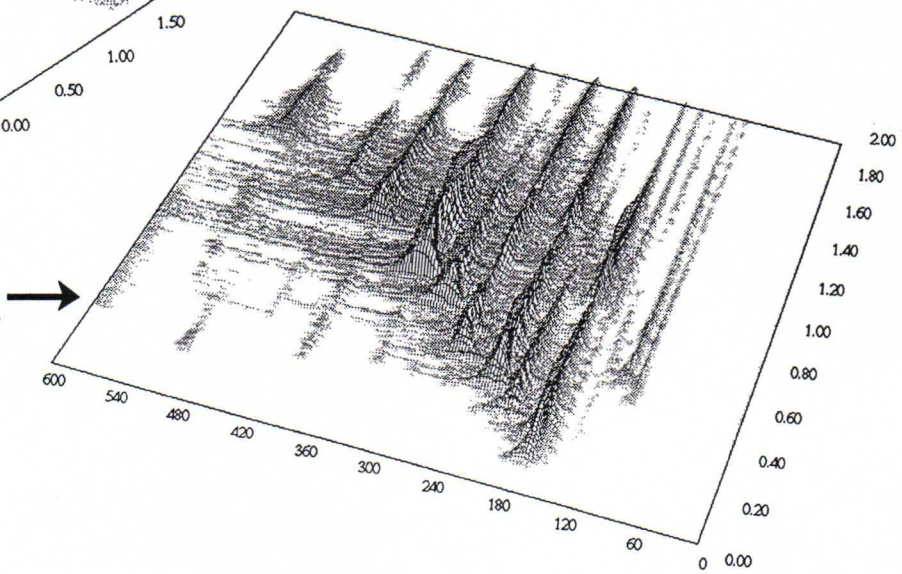


fig.11 : Natural instrumental sound (harp arpeggio) ΔT=0,01sec ΔF=3Hz (200 cells)